ON THE REVELATION PRINCIPLE AND RECIPROCAL MECHANISMS IN COMPETING MECHANISM GAMES

MICHAEL PETERS

Abstract. This paper shows how to characterize the set of outcomes functions that can be supported as equilibrium outcome functions in competing mechanism games. We describe a set of mechanisms we refer to as *reciprocal mechanisms.* It is show that the set of outcome functions supportable as equilibria in a game in which principals offer reciprocal mechanisms is the same as the set of outcomes supportable by a centralized mechaism designer. It is then shown that any outcome function that can be supported as an equilibrium outcome in *any* competing mechanism game can also be supported as an equilibrium outcome function in a game in which all players offer reciprocal mechanisms. In this equilibrium, all players offer the same reciprocal mechanism which implements a collection of direct mechanisms in which players report their payoff types truthfully to other players.

1. Introduction

It has been known for some time that the 'revelation principle' doesn't hold in competing mechanism games. What this means is that modeling competing mechanism designers as if they offered the usual kind of direct mechanisms in which agent reports about their payoff types are converted into actions, will only capture some of the things that can be supported as equilibrium in competing mechanism games. This argument stems from a remark in (McAfee 1993) - since agents observe the mechanisms that were offered by the other mechanism designers, their types should be defined broadly enough to allow them to convey this payoff information. McAfee didn't make anything of this, but the subsequent literature offers many simple examples to illustrate why this actually makes a difference. Examples of equilibrium outcome functions in competing mechanism games that can't be supported as equilibria when designers offer naive direct mechanisms have been given by (Peck 1995), (Epstein and Peters 1999), (Martimort and Stole 2002), and (Peters 2001) among others.

(Epstein and Peters 1999) provides a type space and set of mechanisms which allows agents to convey market information along with information about their payoff type. They show that every equilibrium in a competing mechanism game is equivalent to an equilibrium relative to what they called the universal set of mechanisms. The universal set of mechanisms contains only mechanisms that convert type reports into outcome. In the equilibrium relative to this set of mechanisms, every agent reports his or her type truthfully. However, types are taken from the 'universal set of types', which is broad enough to convey both the agent's payoff type, and his market information.

Very Preliminary Version - January 8, 2011.

The universal set of types illustrates two things. First, the McAfee idea that type should be reinterpreted in competing mechanism games is right, but unusable. Universal types are complex, and have properties that do not lend themselves to applications.[1] However, the approach did illustrate two useful facts. First, equilibria in naive mechanisms are typically robust to expansions of the set of feasible mechanisms. What this means is that the equilibria that McAfee was able to describe - in which competing mechanism designers simple don't ask agents for market information because they aren't allowed to - are also equilibria when mechanism designers are allowed to use much more sophisticated mechanisms. The complication in competing mechanisms is that equilibrium supports allocations that aren't supportable when mechanism designers are restricted to smaller sets of mechanisms.

The literature on common agency (many competing principals, but only a single agent) tried to remedy this by abandoning the revelation principle, and simply asking for some set of indirect mechanisms that could be used to support all outcomes that might quality as common agency equilibrium. (Martimort and Stole 2002) and (Peters 2001) show that every (robust) equilibrium relative to any set of indirect mechanisms in common agency is an equilibrium relative to the set of menus. (Pavan and Calzolari 2009) show a similar result for common agency using what they call the set of 'extended direct mechanisms'. All robust pure equilibrium in common agency are equilibrium relative to the set of extended direct mechanisms.

As useful as the common agency tools are, they have two shortcomings. First, common agency is special since there can only be one agent, and principals can't communicate. Second, though the set of mechanisms (menus) that this literature offers is considerably simpler than the universal set of mechanisms, they are not sufficiently structured to allow a characterization of supportable outcomes.[2]

(Yamashita 2007) has recently suggested a way to extend the common agency logic to problems in which each principal has many agents. As in common agency, principals simply ask agents what to do, and commit themselves to carry out the recommendation as long as the majority of the recommendations agree. A characterization theorem for this case is given by (Peters and Troncoso-Valverde 2009) for competing mechanism games with at least four players and at least one principal. Oddly, a simple common agency does not fit into this environment.

This paper does two things. First, it provides a simpler set of indirect mechanisms which can be used to characterize equilibrium outcomes in all environments. We call these mechanisms reciprocal mechanisms. They play the same role as menus in common agency, in the sense that an outcome function supportable as an equilibrium in some game of competing mechanism designers can also be supported as an equilibrium in reciprocal mechanisms. However, it applies to environments with many agents as well as to environments with no agents at all. In a game in reciprocal mechanisms, players send each other messages of three kinds. The first are public messages that play a dual role, both conveying market information and specifying a commitment. The second kind of message is a private message conveying a player's payoff type. The third is a private correlating message whose role is to support randomization and correlation.

[1] For example, agents types typically depend on the mechanism that a seller offers.

[2] Characterizations of outcomes for special environments have been given by (Peters and Troncoso-Valverde 2009). Though it might not be apparent why yet, we would also include (Tennenholtz 2004) and (A.T. Kalai and Samet 2007) in this category.

The first part of our main theorem is a surprisingly simple characterization theorem. Every outcome function that can be supported by a centralized mechanism designer can be supported as an equilibrium in reciprocal mechanisms, i.e., as an equilibrium in competing mechanisms. We emphasize that we impose no restriction on the number of players. Nor do we focus on or require players to use pure strategies or to enjoy non-random outcomes. Indeed, since a centralized mechanism designer can support correlated outcomes, we can show how to support these as equilibria in a competing mechanism game.

The 'market' messages we add are based on an idea from computer science[3] suggesting that computer programs exhibit a kind of duality. A program is a set of instructions that converts some set of inputs into an output. At the same time, the syntax of the program can can serve as an input into other programs. We are just going to treat contracts like computer programs. Instead of using the syntax of the contracts directly, we are going to create a set of message an associated each contract with one of these messages.[4] We will then simply think of contract offers as public messages and allow players to write contracts that make commitments conditional on these messages.

However, we don't need elaborate mathematics to do this. The reason is that we are simply trying to find a way to characterize all equilibrium outcomes. For this reason we are free to restrict the set of contracts and messages that players can use in much the same way that the revelation principal restricts players to sending type reports instead of arbitrary messages. In fact, the set of feasible contracts is constructed using a method that is similar to that in (Pavan and Calzolari 2008) who exploit the idea of a player's extended type. This is supposed to describe the outcome that the player expects to occur in the game. The set of contracts we use coincident with this set of types.

To see the rough idea, consider a simple symmetric prisoner's dilemma game with actions C and D, and restrict attention for the moment to pure outcomes. Begin by defining a message θ^*, which either player can announce publicly at the first stage of our competing mechanism game. Announcing this message commits the player to the following action which depend on the public message of the other player:

$$\theta^* \equiv \begin{cases} C \text{ if player 2 announces } \theta^* \\ D \text{ otherwise.} \end{cases}$$

Add to this two additional messages, θ^c. Announcing θ^c at the beginning of the game commits the player to use the cooperative action no matter what public message the other player uses, and θ^d which commits to the non-cooperative action no matter what message the other player announces. We then simply define the set of feasible reciprocal mechanisms as $\{\theta^*, \theta^c, \theta^d\}$. Each of the symbols θ^*, θ^c and θ^d plays a dual role as it represents a specific mechanism, but is also used as a possible message which is used by the mechanisms or contract of the other player.

The normal form of the game in reciprocal mechanisms is simply the following (replacing payoffs with outcomes):

[3]Von Neumann

[4]This is similar to the approach in (Peters and Szentes 2008) where the literal syntax of contracts is re-encoded into a integer.

	θ^*	θ^c	θ^d
θ^*	(CC)	(DC)	(DD)
θ^c	(CD)	(CC)	(CD)
θ^d	(DD)	(DC)	(DD)

It is immediate that in this new normal form in reciprocal mechanisms, there is an equilibrium in which both players publicly announce the mechanism θ^*. In this equilibrium cooperation occurs. Evidently (DD) can also be supported as an equilibrium.

It is also straightforward to extend this argument to arbitrary (finite) complete information games. Suppose there are n players, each of whom has a finite set of feasible actions A_i. Define $A = \prod_{i=1,\ldots n} A_i$ and $A_{-i} = \prod_{j\neq i} A_j$. Again, we will stick for the moment to pure strategies, and define a collection of constant mechanisms for each player. To make things a little simpler, suppose each player has k pure actions.[5] Then create k messages $\theta^i_{a_1}$ to $\theta^i_{a_k}$ for player i. Announcing any of these messages publicly commits i to play the corresponding action no matter what messages other players announce.

Now we can create a set of reciprocal mechanisms which will be denoted by messages like $\theta^i_{\left(a_1,a'_i\right),\ldots,\left(a_n,a'_n\right)}$. The objects a_i are elements of the feasible set of actions for player i. One way to think of these is that they are the cooperative actions that players are trying to support with their reciprocal mechanisms. The object a'_i is a vector of $n-1$ actions from i's action set. Informally, these represent punishments that i will implement against each of the other players when they refuse to cooperate. So the index (a_i, a'_i) has n components. There are k^{n^2} reciprocal mechanisms like this. The set of mechanisms available to each player i will then be the set of constant mechanisms plus this set of reciprocal mechanisms.

The public message $\theta_{\left(a_1,a'_1\right),\ldots\left(a_i,a'_i\right),\ldots,\left(a_n,a'_n\right)}$ commits player i as follows:

$$\theta_{\left(a_1,a'_1\right),\ldots\left(a_i,a'_i\right),\ldots,\left(a_n,a'_n\right)}\left(\theta'_{-i}\right) =$$

$$\begin{cases} a'_{ij} & \exists! j : \theta'_j \neq \theta_{\left(a_1,a'_{-1}\right),\ldots,\left(a_i,a'_{-i}\right),\ldots,\left(a_n,a'_n\right)} \\ a_i & \text{otherwise.} \end{cases} \tag{1.1}$$

In the notation above $\exists!$ means "there exists a unique".

Each mechanism like this has a cooperative action a_i and an array of punishments $a'_i = \{a'_{i1}, \ldots, a'_{im}\}$ for each of the other players. If all the other players offer the same reciprocal mechanism as player i, then i will respond with his cooperative action a_i. If all but one of the players offer the same reciprocal mechanism as player i, then i will respond with the punishment action a'_{ij} for that player. It is immediate that this set of reciprocal mechanisms supports every action profile in which each player receives at least his minmax payoff as a Nash Equilibrium in mechanisms.

This illustrates how reciprocal mechanisms work. We now extend the idea to allow for randomization and incomplete information.

[5]This is without loss of generality since we can modify payoff functions to assign low payoffs when i takes actions that aren't feasible in the original game.

2. Incomplete Information Games

In a game of incomplete information, there are n players. Each player has a finite action set A_i and a finite type set T. In standard notation A, A_{-i} represent cross product spaces representing all players actions and the actions of all the players other than i, respectively. Types are jointly distributed on T^m according to some common prior and preferences of player i are given by $u_i : A \times T \to \mathbb{R}$. Players have expected utility preferences over lotteries. Note that because the distribution of types is arbitrary, it is without loss of generality to assume that all players types lie in the same space. We are also going to assume that the set of potential types is at least as large as the set of actions available to each player.

Let Q_i, Q and Q_{-i} represent mixtures over A_i, A and A_{-i} respectively. Since players have expected utility preferences we write $u(Q,t)$ to be the expected utility associated with the mixture Q when types are i. An *outcome function* is a mapping $\omega : T^n \to Q$. An outcome function is *incentive compatible if*

$$\mathbb{E}\left\{u\left(\omega\left(t\right),t\right)|t_i\right\} \geq \mathbb{E}\left\{u\left(\omega\left(t_i',t_{-i}\right),t\right)|t_i\right\}. \tag{2.1}$$

A punishment $\rho_i : T_{-i} \to P_{-i}$ is used when player i chooses not to participate in the mechanism that implements ω. The outcome function is *individually rational* if there is a punishment ρ_i for each player i such that for every i

$$\mathbb{E}\left\{u\left(\omega\left(t\right),t\right)|t_i\right\} \geq \max_{a_i \in A_i} \mathbb{E}\left\{u\left(a_i,\rho_i\left(t_{-i}\right)\right),t|t_i\right\}. \tag{2.2}$$

Players choose mechanisms that commit them to actions that depend on messages that they send and receive. Messages available to player i are elements of a measurable space $\mathcal{M}_i$. Each $m \in \mathcal{M}_i$ is a sequence of messages that can be eithe public or private. Let ι_j^i be a measure on $\mathcal{M}_j$ describing which messages of player j are observable by i. Each player i is assumed to be able to observe all of his own messages. All observable messages are assumed to be contractable.

We model mechanisms by assuming that the sequence m always includes messages that describe player i's commitments. Formally $\mathcal{M}_i$ consists of a set of equivalence classes $\mu_i(m)$ such that each message $m' \in \mu_i(m)$ is associated with the same mapping $\gamma_i : \prod_j \mathcal{M}_i \to A_i$. This mapping is assumed to be measurable with respect to i's information $\left\{\iota_j^i\right\}_{j\neq i}$. This mapping describes how player i's action is determined by the messages that he observes. The set of feasible commitments for player i is a set Γ_i - each mechanism in this set being associated with a unique equivalence class.

A message m is said to represent an *instrutable mechanism* if the mechanism γ_i associated with $\mu_i(m)$ is an onto function from $\mu_i(m)$ to A_i, and if the only subset of $\mu_i(m)$ that is measurable with respect to any ι_i^j is $\mu_i(m)$ itself. We assume throughout that the set of mechanisms available to every player includes an inscrutable mechanism.

The objects Γ , $\{\mathcal{M}_i\}_{i=1,\dots,n}$ and $\left\{\iota_j^i\right\}_{i,j=1,\dots n}$ represent a particular model of a competing mechanism game. The mechanisms described by each of the $\mathcal{M}_i$ are analogous indirect mechanism in the usual problem. The description is intended to capture a broad variety of different models. For example, messages, including commitments, may be either public or private. They can be made sequentially or simultaneously. Particular models of competing mechanisms may also include restrictions on mechanism designers' ability to contract. In a typical common agency model, for example, mechanisms are annouced publicly, but no player is allowed to condition directly on these publicly verifiable messages. So the set of feasible mechansims may consist solely of mechanisms that are measurable with respect to measures that are coarser than the information measures given by $\{\mu_j^i\}$.

The formulation here simply assumes there is a set of players, all of whom could potentially offer contracts. The special models that have been examined in the literature are special cases. For example the principle agent formulation ((Epstein and Peters 1999), (Yamashita 2007), (Han 2006)) assumes that players can be divided into two groups consisting of uninformed principals who possess the ability to write contracts, and informed agents who can't write contracts at all. Formally, we simply endow 'agents' with the ability to condition only on their own messages. Common agency (Peters 2001) or (Martimort and Stole 1998) imposes the additional restriction that there is only a single agent, and prevents the 'principals' from contracting with each other.

It isn't at all obvious what the right way to model indirect mechanisms is in the competing mechanism game. Menu mechanisms in common agency, or competition in reserve prices among aucitoneers ((Peters and Severinov 1997)) simply impose restrictions on what mechanism designers can do. Complex issues associated with contractual robustness, and the infinite regress associated with allowing contracts to depend on other contracts are addressed in (Epstein and Peters 1999) and (Peters and Szentes 2008). Here we are simply going to assume that the modeller has chosen methods to resolve these issues and included them in the specification of the set of messages.

Given this specification, the equilibrium in the competing mechanism game is defined in the usual way. A strategy rule specifies for each player and for each of his possible types, a randomization over the set of mechanisms and rules that specify randomizations over the set of messages sent to other players conditional whatever information is available at the time the message is sent. We focus on Bayesian equilibrium (as does the revelation principle) because we are interested in characterizing all possible outcomes, and all the refined equilibrium are also Bayesian equilibrium.

2.1. **Reciprocal Mechanisms.** The objective here is to show that there is a relatively simple competing mechanism game that can be used to understand equilibria in all competing mechanism games. This game takes place in two stages. At the first stage each player simultaneous announces a public message that describes his commitments in the game. In the second round each player privately sends a message in $[T \times [0,1]]$ to each of the other players. The first element of this message is a type report, the second is a correlating message used to support randomization. In the formalism of the previous section the message space consists of a sequence of two signals, the first from Θ_i are all public, the second from $(T \times [0,1])^n$ are all private.

The public messages in Θ_i are all tied to commitments which may be of two possible types. We describe the messages then associate a mechanism with each of them. First, we focus on a subset of the set of measurable mappings c_i : $(T \times [0,1])^n \times (T \times [0,1])^{n-1} \to A_i$. The arguments of this function are the player's own messages, and the messages that he receives privately from the other players.

A mechanism d_i is referred to as a *direct mechanism* if it is a measurable function from $(T \times [0,1])^n \times (T \times [0,1])^{n-1} \to A_i$, if d_i implements some arbitrary fixed action unless the first n messages are all the same, and if it depends only on the fractional part of the sum of the n correlating messages. We explain how these correlating messages are used below. The set of direct mechanisms available to player i is D_i. A *punishment mechanism* p_i^j for player i to use against player j is a direct mechanism that is independent of the messages of player j. Let D_i be the set of direct mechanisms available to player i. We let D_i refer to the set of direct mechanisms.

The second kind of commitment is a reciprocal mechanism. To construct these, we begin with the public messages that represent them. Let $\delta_i = \left\{ d_i, \left\{ p_i^j \right\}_{j \neq i} \right\}$ consist of a direct mechanism and a list of punishment mechanisms. This list consists of a different punishment mechanism for each of the other players. The notation $\delta = \{\delta_1, \ldots, \delta_m\}$ represents a list of such lists, while Δ is the set of all δ. We are going to index the set of feasible mechanism for player i by $\Theta_i = \Delta \cup D_i$. If $\theta_i \in D_i$, then θ_i commits player i to the corresponding direct mechanism.

We need to associate a unique commitment with each element of Δ. Use the notation θ_i^δ to be the mechanism associated with the message δ, and let d_i^δ to refer the direct mechanism associated with the i^{th} element of the list δ. The *reciprocal mechanism* θ_i^δ given by the mapping

$$\theta_i^\delta \left(\theta'\right) = \begin{cases} p_i^j & \exists ! j : \theta'_j \neq \theta_j^\delta \\ d_i^\delta & \text{otherwise.} \end{cases} \tag{2.3}$$

The notation $\exists!$ means "there exists a unique". This commits player i to the direct mechanism d_i^δ (from the list δ) unless there is a unique player j who fails to use the mechanism θ_j^δ. In that case, the mechanism commits i to use the punishment mechanism p_i^j designed for that player.

3. Theorem

The set of Θ_i of mechanisms available to each player is small. One can imagine mechanisms more complicated than direct mechanisms. Reciprocal mechanism especially are restrictive. They allow players to respond to deviations, but only in a manner that is independent of what the deviation is. Nonetheless, we can show that when players are restricted to choose their mechanism in Θ_i, any incentive compatible and individually rational outcome function can be supported as an equilibrium in the corresponding competing mechanism game.

Let ω be an outcome function that is incentive compatible and individually rational in the sense of (2.1) and (2.2). Recall that an outcome function has the property that for every array of types t, $\omega\left(t\right)$is some joint distribution on A.

Theorem 1. *There is a Bayesian equilibrium in the contracting game with reciprocal mechanisms Θ that supports the outcome function ω if and only if ω is implementable by a mechanism designer in the sense that is satisfies (2.1) and (2.2). In this equilibrium, all players use a common mechanism θ^δ, players report their types truthfully at the second round to each player, and each player chooses a correlating message uniformly from the interval* $[0, 1]$.

Proof. Index the action profiles in A in some arbitrary way. Let $\omega^k\left(t\right)$ be the probability assigned to action profile a^k by the outcome function ω when player types are given by the vector t. The notation a_i^k means the action taken by player i in action profile a^k. As mentioned above, we restrict attention to mechanisms where i constrains himself to send the same message to all the other players. Let $x = (x_i, x_{-i})$ be the vector of signals sent by all the players. Now define

$$d_i^{\omega}(t,x) = \left\{ a_i^k : k = \min_{k'} : \sum_{\tau=1}^{k'} \omega^{\tau}(t) \le \lfloor \sum_j x_j \rfloor \right\}. \tag{3.1}$$

The notation $\lfloor y \rfloor$ means the fractional part of the real number y. This functions aggregates the signals x into a number between 0 and 1, uses this to choose the index of the action profile in A that depends on the type reports of all the players, then directs player i to take his part in this action profile.

To see how this is going to implement the desired outcome, suppose first that each of the players chooses x_i using a uniform distribution. The random variable $\lfloor \sum_j x_j \rfloor$ is uniformly distributed on $[0,1]$. In that case, the function $\{d_1^{\omega}(t,x), \ldots, d_n^{\omega}(t,x)\}$ implements the action profile a^k with probability $\omega^k(t)$ as the outcome function requires. The more interesting property of this construction is that for each value of x_i, $\lfloor x_i + \sum_{j\neq i} x_j \rfloor$ is also uniformly distributed on $[0,1]$.[6] This means that whatever else is happening, it is a best reply for each player i to select a number x_i using a uniform distribution provided he believes the others are doing the same thing.

We use exactly the same construction for the punishment

$$p_{ij}^{\rho_j}(t,x) = \left\{ a_i^k : k = \min_{k'} : \sum_{\tau=1}^{k'} \rho_j^{\tau}(t) \le \lfloor \sum_{\tau\neq j} x_j \rfloor \right\}. \tag{3.2}$$

Let $\delta_i^{\omega} = \left\{ d_i^{\omega}, \{p_{ij}^{\rho_j}\}_{j\neq i} \right\}$ where $d_i^{\omega}$and $p_{ij}^{\rho_j}$ are defined by (3.1) and (3.2), and $\delta^{\omega} = \{d_i^{\omega}\}_{i=1,\ldots,n}$. We now claim that if the outcome function ω is implementable, then there is a Bayesian equilibrium in the competing mechanism game in which each player i offers reciprocal mechanism $\theta_i^{\delta^{\omega}}$ as defined by (2.3) using the list δ_i^{ω}.

To see why, imagine first that all players offer this reciprocal mechanism. Then all players are constrained to use the function d_i^{ω} to translate messages into actions. We have already explained that whatever type report a player sends to other players, he is completely indifferent about the message that he sends as long as he believes the others messages are chosen uniformly. As a consequence, there is a continuation equilibrium in which each player chooses his message uniformly. If all other players are revealing their true types to other players, (recall we are assuming here that every player is using a direct mechanism which constrains each of them to send the same message to every other player), then the payoff to player i if he also reveals his true type is

$$\begin{aligned}\mathbb{E}\{u(d_1^{\omega}(t,x), \ldots, d_m^{\omega}(t,x), t) | t_i\} &= \mathbb{E}\{u(\omega(t_i, t_{-i}), t) | t_i\} \\ &\ge \mathbb{E}\{u(\omega(t_i', t_{-i}), t) | t_i\}\end{aligned}$$

where this last expression is the payoff he gets if he lies about his type. The inequality follows from (2.3) and (2.1).

Now suppose that i deviates to some alternative contract c'. This must be unprofitable. To ensure this, we need to construct continuation play that makes i worse, conditional on the fact that, as a unilateral deviator, the other players are

[6]This device is from the paper (A.T. Kalai and Samet 2007). A proof of this last property is given in (Peters and Troncoso-Valverde 2009). This proof also shows why $\lfloor \sum_j x_j \rfloor$ must be uniformly distributed.

committed by the reciprocal contract, to carry out the punishments $p^{\rho_i}_{-i,i}$ against i. We assume that each player j continues to report type truthfully when implementing punishments, and that whenever i's contract asks the others to make reports, they report some arbitrary messages $\left(\bar{t}_{-i}, \overline{x}_{-i}\right)$. Then since the action that player i takes cannot depend on the types of the other players, the payoff when player i of type t_i deviates is

$$\max_{\{t_{ij},x_{ij}\}_{j\neq i}} \mathbb{E}\left\{u\left(c'\left(\{t_{ij},x_{ij}\}_{j\neq i},\left(\bar{t}_{-i},\overline{x}_{-i}\right)\right),p^{\rho_i}_{-i,i}\left(t_{-i},x_{-i}\right),t\right)|t_i\right\} \leq$$

$$\max_{a_i}\mathbb{E}\left\{u\left(a_i,p^{\rho_i}_{-i,i}\left(t_{-i},x_{-i}\right),t\right)|t_i\right\} \leq$$

$$\mathbb{E}\left\{u\left(\omega\left(t_i,t_{-i}\right),t\right)|t_i\right\}$$

where the last line follows from (2.2).

To prove the other direction, begin with an outcome function ω that is supportable as a Bayesian equilibrium in reciprocal mechanisms. Player i's equilibrium payoff is

$$\mathbb{E}\left\{u\left(\omega\left(t\right),t\right)|t_i\right\}. \tag{3.3}$$

Let c' be any contract for which i's action is an onto function only his own correlating message x_i. In any equilibrium relative to reciprocal mechanism, this deviation must be unprofitable. In response to such a deviation, the other players will respond with some randomization over punishment mechanisms $\left\{p^{\rho_j}_{ij}\right\}_{j\neq i}$. The randomization is possible because players may be mixing when choosing among mechanisms. The messages and punishments that the others use can't depend on i's choice of action. Furthermore, choosing a payoff maximizing message at the second stage is equivalent to choosing a payoff maximizing action at the second stage since the contract is onto. It then follows

$$\mathbb{E}\left\{\max_{t_i'}\mathbb{E}_{\left(t'_{-i},t_{-i},x_{-i}\right)}u\left(a_i^{t_i'},\tilde{p}_{-i}\left(t'_{-i},x_{-i}\right),\left(t_i,t_{-i}\right)\right)|t_i\right\} \leq$$

$$\mathbb{E}\left\{u\left(\omega\left(t\right),t\right)|t_i\right\}.$$

This is almost the expression we want, except that when taking expectations on the left hand side. To complete the argument, let $\rho'\left(t_{-i}\right)$ be the probability distribution over the actions A_{-i} conditional on t_{-i} that is induced by the equilibrium strategies of the players other than i in the continuation equilibrium. Then from the inequality above

$$\max_{a_i}\mathbb{E}\left\{u\left(a_i,\rho'\left(t_{-i}\right),\left(t_i,t_{-i}\right)\right)|t_i\right\} \leq \mathbb{E}\left\{u\left(\omega\left(t\right),t\right)|t_i\right\},$$

which completes the proof. □

4. Equilibrium in Inscrutable Indirect Mechanisms

Theorem 1 provides a characterization of outcome functions supportable in reciprocal mechanism games. Reciprocal mechanism are restrictive as we have mentioned. For these reasons, it may be more sensible to model competing mechanism games with a more realistic looking set of indirect mechanisms. The next theorem simply states that if this set of indirect mechanisms satisfies the inscrutability assumption we described above, then any outcome function that can be supported as an equilibrium relative that set of mechanisms can also be supported as an equilibrium in reciprocal mechanisms.

All that is required to to this is to show that if an outcome function is supportable as equilibrium relative to an inscrutable set of indirect mechanisms, then it satisfies (2.2).

Theorem 2. *If an outcome function ω is supportable as an equilibrium in inscrutable indirect mechanisms, then it is supportable as an equilibrium in reciprocal contracts.*

Proof. The proof mimics the second part of the proof in Theorem 1. Let $\omega(t)$ be the outcome function supported by some equilibrium in inscrutable indirect mechanisms. Player i's payoff is given by (3.3). Since player i can Let c' be a deviation to an inscrutable contract in which i's action is an onto function from $\mathcal{M}_i^{[0,1]}$ to A_i. Let $\rho'(t_{-i})$ be the type contingent randomization over A_{-i} associated with the continuation equilibrium associated with this deviation. Since choosing a message in $\mathcal{M}_i^{[0,1]}$ is equivalent to choosing an action, it follows from the fact that the deviation is unprofitable that

$$\max_{a_i} \mathbb{E}\left\{u\left(a_i, \rho'\left(t_{-i}\right), \left(t_i, t_{-i}\right)\right)|t_i\right\} \leq \mathbb{E}\left\{u\left(\omega\left(t\right), t\right)|t_i\right\},$$

□

Theorem 2 differs slightly from theorem 1 since it only goes in one direction. Not all the outcome functions that can be supported with reciprocal mechanisms can be supported as equilibrium relative to some arbitrary set of indirect mechanisms.

5. An Example

To illustrate, we consider a well known issue in mechanism design. Suppose there are two sellers and two buyers (i.e. four players in all). Each seller has a single unit of output to sell to which he or she assigns a value of v_m. Each buyer has a private valuation, either v_l or v_h ranked the obvious way with $v_l < v_m < v_h$. Payoffs to the seller are equal to the money he receives less his value if he trades, payoffs to each buyer are equal to their private valuation when they succeed in trading, less the money they pay. Each seller offers a mechanism, each buyer chooses to participate in one and only one mechanism. We assume that valuations are correlated. To make life simple suppose that both valuations are the same with probability $q > \frac{1}{2}$ and that they are equally likely to (both be) v_h or v_l in that case.

As for feasible actions, each seller can choose to give his good to either of the two buyers, or to keep it. He can also choose to make transfers to either or both buyers. We will simply ignore the fact that the set of feasible transfers isn't finite. Similarly, buyers can offer to trade with either of the sellers or not to trade at all. They can also make transfers to either or both sellers.

First of all, we describe the analog of the (Cremer and McLean 1988) result for this environment. We would like to have an allocation rule that is ex post efficient in the sense that a buyer trades with some seller if and only if his valuation exceeds v_m. Furthermore, we want the interim expected payoff of each buyer to be zero. Notice that we can accomplish this in a centralized mechanism because the buyer who has type v_l believes that the other buyer has type v_l with probability q, while the buyer with type v_h believes the other buyer has type v_l with probability $1-q$.

Then assuming we have the high type buyer trade with one of the sellers at price v_m, we need to design a pair of transfers (p_l, p_h) so that

$$v_h - v_m + qp_h + (1-q)\, p_l = 0$$

and

$$(1-q)\, p_h + qp_l = 0.$$

Since $q > \frac{1}{2}$ and $v_h > v_m$, this pair exists and is unique. It is straightforward that $p_l > 0$, and $p_h < 0$. Adding the two equations together gives $(v_h - v_m) + p_h + p_l = 0$

Our centralized mechanism has each buyer report his or her type. A buyer who reports v_h is given a unit of output and pays a price v_m. He also pays a fee p_h if the other buyer's reported type is v_h, and receives the transfer p_l if the other buyer's reported type is v_l. A buyer who reports a type v_l doesn't get any output, but still pays the fee or receives the transfer depending on the reported type of the other buyer. It is straightforward that a high type buyer doesn't want to misreport because he loses the good but still ends up paying a positive amount. Similarly for the low type.

The surplus earned by sellers under this scheme is $(v_h - v_m)/2$, which is the surplus sellers earn with complete information. To explain why the mechanism is individually rational, we need to explain what happens when one of the players unilaterally decides not to participate. Mechanism design as it is described in (2.2) assumes that participating players can impose arbitrary punishments. To start, assume we impose the obvious punishment that participants refuse to trade with non-participants. Then refusing to participate yields every player a payoff 0 and the mechanism is individually rational as in (2.2).

An obvious use of competing mechanisms is to try to explain away this odd result by using the argument that it can't be an equilibrium for competing sellers to leave both buyers with zero expected surplus. However, we can support such an outcome with reciprocal mechanisms. Each seller's mechanism takes input from the other seller and both buyers. We describe a set of direct mechanisms that implement the outcome. As no one randomizes in this outcome, we don't need signals at all. Seller i uses mechanism d_i^s which commits to sell to buyer i at a price v_m if buyer i reports a high valuation, otherwise seller i keeps his good. Seller i commits to pay buyer i a transfer p_l if buyer j reports type v_l. Buyer i uses mechanism d_i^b which commits to buy from seller i and pay a price v_m if buyer j reports type v_l, and to pay seller i $v_m - p_h$ otherwise (recall that $p_h < 0$). Recall that these target mechanisms commit the buyers to send the same messages to everyone.

To complete the description of the corresponding reciprocal mechanisms, we need to describe the punishments that go along with these direct mechanisms. This is straightforward here, all players simply commit not to trade when there is a deviation. Recall these are commitments built into contracts, not presumptions about how players will play a continuation game. The punishment p^* commits to no trade, zero transfers for all messages. So we have the reciprocal mechanisms

$$\theta^j_{\left\{\left(d_i^s, p^*\right), \left(d_i^b, p^*\right)\right\}_{i=1,2}} = \begin{cases} p^*\left(v_1, v_2\right) & \exists! k : \theta'_k \neq \theta^k_{\left\{\left(d_i^s, p^*\right), \left(d_i^b, p^*\right)\right\}_{i=1,2}} \\ d_j\left(v_1, v_2\right) & \text{otherwise.} \end{cases}$$

6. Literature

We have shown that all equilibria of competing mechanism games can be understood using reciprocal mechanisms. The advantage of this is that reciprocal mechanisms are conceptually no more difficult to work with than ordinary direct mechanisms. So reciprocal mechanisms provide a useful analytic approach for problems in which a broad class of mechanisms is feasible.

One consequence of this observation is that the set of allocations that can be supported as equilibrium with competing mechanisms is large. This fact has been observed before. Starting with the large literature on delegation games ((Fershtman and Judd 1987, Fershtman and Kalai 1997)), a number of papers have shown large equilibrium sets for special cases ((Katz 2006, Tennenholtz 2004, Yamashita 2007, Peters and Troncoso-Valverde 2009)). Our paper differs from these in two ways. First we impose no restrictions on the environment. (Katz 2006, Tennenholtz 2004), for example, assume complete information. (Yamashita 2007) assumes that players who offer contracts have no private information, and restricts the number of players. (Peters and Troncoso-Valverde 2009) restricts the number of players. We impose none of these restrictions.

Secondly, like the papers by (A.T. Kalai and Samet 2007)[7] and (Peters and Szentes 2008) we provide a complete characterization of supportable equilibrium outcomes rather than simply illustrating that a large number of equilibrium outcomes can be supported. However we do not assume, as do (A.T. Kalai and Samet 2007) that players have complete information. We do not restrict the set of feasible mechanism nor the set of equilibrium outcomes as in (Peters and Szentes 2008).

Like direct mechanisms, reciprocal mechanisms make it possible to understand equilibrium outcomes with competition without worrying about the intricacies of particular indirect mechanisms that are used in practice. Apart from the standard logic of incentive constraints, reciprocal mechanisms simply add the logic that if everyone else wants to do something, it is simple to write a contract that commits you to do it too.

References

A.T. Kalai, E. Kalai, E. L., and D. Samet (2007): "A Commitment Folk Theorem," Working paper, Tel-Aviv University.

Cremer, J., and R. McLean (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Economtrica*, 56(6), 1247–1259.

Epstein, L., and M. Peters (1999): "A Revelation Principle for Competing Mechanisms," *Journal of Economic Theory*, 88(1), 119–161.

Fershtman, C., and K. Judd (1987): "Equilibrium Incentives in Oligopoly," *American Economic Review*, 77, 927–940.

Fershtman, C., and E. Kalai (1997): "Unobserved Delegation," *International Economic Review*, 38(4), 763–74.

Han, S. (2006): "Menu theorems for bilateral contracting," *Journal of Economic Theory*, 127(1), 157–178, available at http://ideas.repec.org/a/eee/jetheo/v131y2006i1p157-178.html.

Katz, M. (2006): "Observable Contracts as Commitments: Interdependent Contracts and Moral Hazard," *Journal of Economics and Management Strategy*, 15(3), 685–706.

Martimort, D., and L. Stole (1998): "Communications Spaces, Equilibria Sets and the Revelation Principle under Common Agency," University of Chicago unublished manuscript.

Martimort, D., and L. Stole (2002): "The Revelation and Taxation Principles in Common Agency Games," *Econometrica*, 70(4), 1659–1673.

[7]We borrowed the randomizing trick in (3.1) from this paper.

McAfee, P. (1993): "Mechanism Design by Competing Sellers," *Econometrica*, 61(6), 1281–1312.

Pavan, A., and G. Calzolari (2008): "On the use of Menus in Sequential Common Agency," *Games and Economic Behavior*, 64, 329–334.

——— (2009): "Sequential Contracting with Multiple Principals," *Journal of Economic Theory*, 50, 503–531.

Peck, J. (1995): "Competing Mechanisms and the Revelation Principle," *Ohio State University Manuscript*, p. 12.

Peters, M. (2001): "Common Agency and the Revelation Principle," *Econometrica*, 69(5), 1349–1372.

Peters, M., and S. Severinov (1997): "Competition Among Sellers who offer Auctions Instead of Prices," *Journal of Economic Theory*, 75(1), 141–179.

Peters, M., and B. Szentes (2008): "Definable and Contractible Contracts," Working paper, University of British Columbia.

Peters, M., and C. Troncoso-Valverde (2009): "A Folk Theorem for Competing Mechanisms," manuscript, University of British Columbia.

Tennenholtz, M. (2004): "Program Equilibrium," *Games and Economic Behavior*, 49(2), 363–373.

Yamashita, T. (2007): "A Revelation Principle and A Folk Theorem without Repitition in Games with Multiple Principles and Agents," manuscript, Stanford University.